(12) **UK Patent Application** (19) **GB** (11) **2 373 907** (13) **A**

(43) Date of A Publication 02.10.2002

(21) Application No 0107931.8

(22) Date of Filing 29.03.2001

(71) Applicant(s)
NEC Technologies (UK) Ltd
(Incorporated in the United Kingdom)
Level 3, The Imperium, Imperial Way, READING,
Berks, RG2 0TD, United Kingdom

(72) Inventor(s)
John Parker

(74) Agent and/or Address for Service
Reddie & Grose
16 Theobalds Road, LONDON, WC1X 8PL,
United Kingdom

(51) INT CL[7]
G06F 3/023

(52) UK CL (Edition T)
G4H HKK H13D H14A

(56) Documents Cited
EP 0924594 A2   US 6005495 A
US 4737980 A

(58) Field of Search
UK CL (Edition S) G4H HKK HKV
INT CL[7] G06F
Online: WPI, EPODOC, PAJ

(54) Abstract Title
Predictive text algorithm

(57) Character selection and display for use in electronic devices such as mobile telephones, from a reduced keypad in which each key has a plurality of characters assigned to it. The character selection method comprises the steps of: detecting a key press; selecting one of the characters assigned to the key for display in dependence on the probability of that character appearing at that location in a character string; repeatedly selecting a further character assigned to the key for display if the first displayed character is not a desired character for that location, until the desired character is displayed; and, assigning the selected character to that location. Preferably, if the character is not positioned within the first three characters of a string, a dictionary is used to search for the entered stem of the word, and to assess the probability of each letter associated with the key being required from the word stem, and display the most likely character as before.
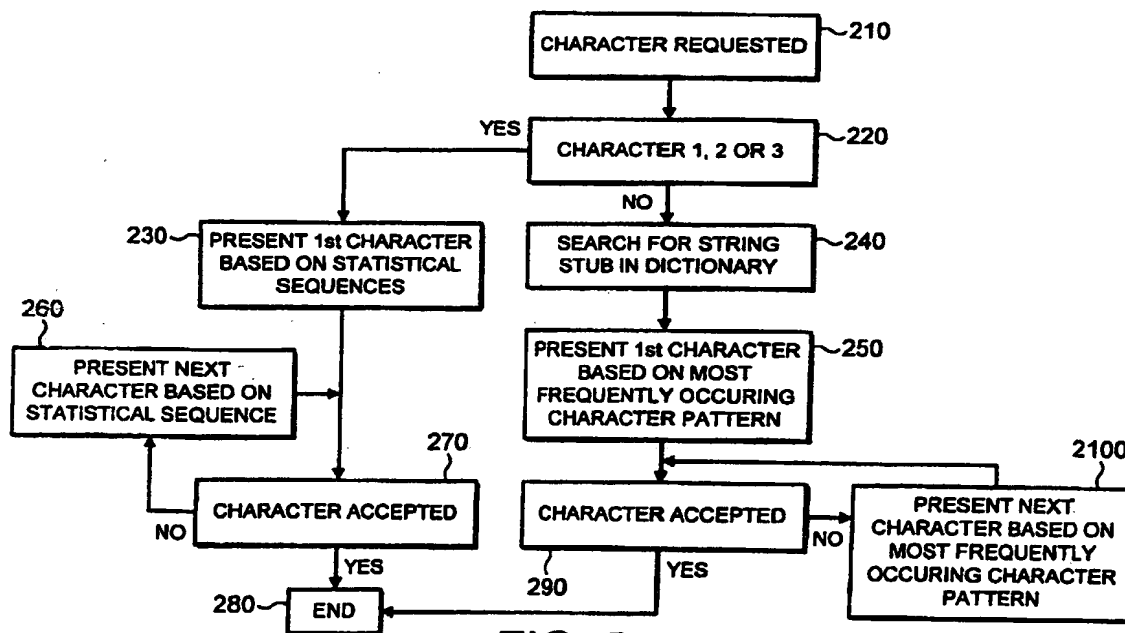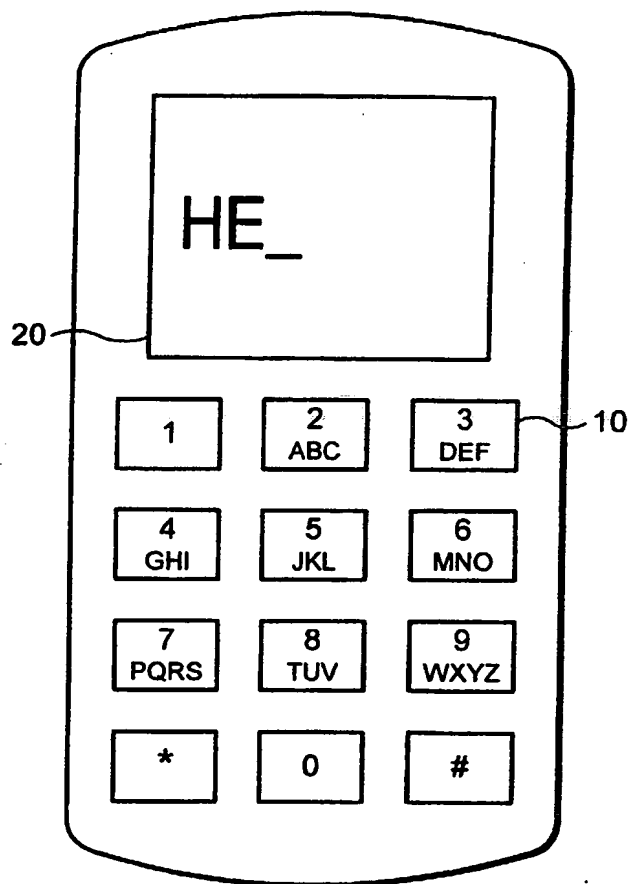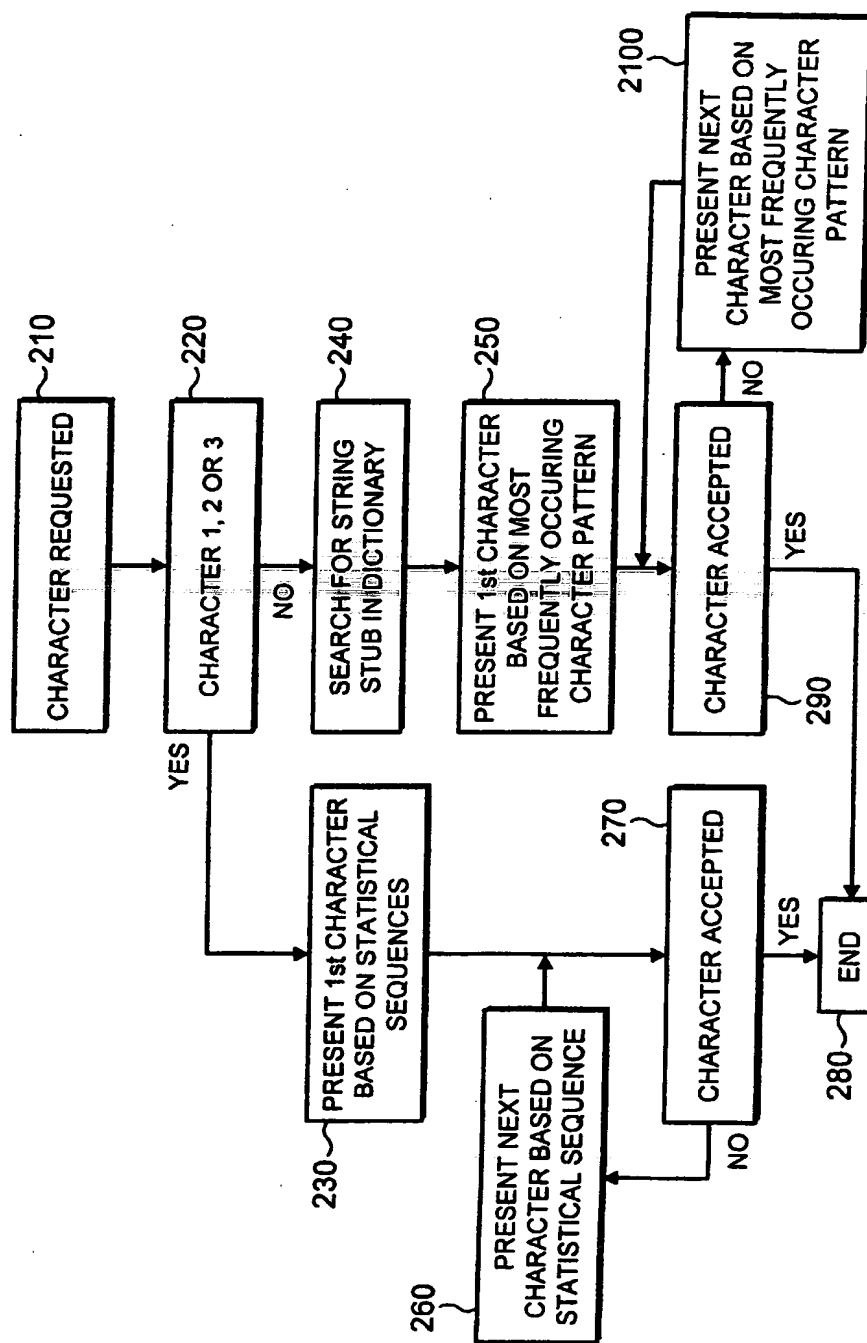
FIG. 2

GB 2 373 907 A

FIG. 1

FIG. 2

CHARACTER REQUESTED ~210

CHARACTER 1, 2 OR 3 ~220

SEARCH FOR STRING STUB IN DICTIONARY ~240

PRESENT 1st CHARACTER BASED ON MOST FREQUENTLY OCCURING CHARACTER PATTERN ~250

PRESENT NEXT CHARACTER BASED ON MOST FREQUENTLY OCCURING CHARACTER PATTERN ~2100

CHARACTER ACCEPTED ~290

PRESENT 1st CHARACTER BASED ON STATISTICAL SEQUENCES ~230

PRESENT NEXT CHARACTER BASED ON STATISTICAL SEQUENCE ~260

CHARACTER ACCEPTED ~270

END ~280

YES

NO

NO

YES

NO

YES

Predictive text algorithm

This invention relates to a method and apparatus for
character selection during string construction such as may
be used in reduced keypad applications e.g. mobile phones.

5   Text messaging on telecommunications devices is now widely
available and phone operators have experienced a huge
increase in the use of text messaging over recent years.
Users require an easy to use service that they can operate
quickly and conveniently. In most systems, each key of
10   the phone unit is mapped to more than one character.
Traditionally there has been no predictive text feature
and selection of the desired character is achieved by
pressing a key more than once while the characters
assigned to that key cycle through a predefined sequence
15   while displayed on the screen. This method necessitates a
significant increase in the number of key presses required
to enter a desired character sequence and is time
consuming to the user.

An algorithm has been proposed in EP-A-0842463 which
20   enables the user to construct a string without the
inconvenience of pressing a key more than once. When the
user requires a character he presses the key associated
with that character only once. At this point the user has
not defined which character is assigned to that position
25   in the string, only that one of the characters associated
with the key is assigned to that position. For any number
of characters entered in a string the algorithm searches
its database for possible words which are constructed from
the key sequence entered by the user. The most likely
30   word is presented to the user on the display screen based
on statistical probability. As more characters are added
to the word the probability of the user requiring a
particular word may change and therefore the word may

change between characters associated with the keys.
Having pressed the keys required to complete the word the
algorithm will recommend the most probable word based on
statistics but will also offer alternative words from its
5   database that are constructed from the same sequence of
key presses.

Problems associated with this algorithm include the large
size of the dictionary database required to contain words
of all lengths. The change in the sequence of characters
10 in the word as more characters are added can be confusing
to the user. Also, text messaging is usually an informal
means of communication with users often using slang
expressions or words not appearing in the dictionary. If
the user is attempting to type a word which is not
15 recognised by the database of the algorithm, that word may
not be offered to the user on completion of the key
presses thus causing confusion and further wasting time.
Since text messaging is also generally a rushed exercise
spelling mistakes are frequent and unimportant and these
20 also cause problems when using the algorithm due to the
reasons mentioned above.

A preferred embodiment of the present invention provides a
predictive character algorithm. The mapping of characters
to a particular key remains unchanged. However, the order
25 in which a character is presented to the user is dependent
upon the preceding characters in the string. The most
likely character is presented first on the screen as
calculated from the statistical database. The statistical
database is generated by considering the probability of a
30 pattern of characters occurring from the beginning of a
word. A further press of the key will present the second
most likely character and so on.

A further preferred embodiment contains two databases; a first database is accessed to assess the likelihood of the user requiring a character associated with a given key for any of, e.g. the first three, letters of a word (as in the first preferred embodiment); the second database is accessed to predict the likelihood of the user requiring a character associated with a given key when the word exceeds a defined number of characters (e.g. 3), the statistical probability for successive characters (e.g. 4 and above) is calculated by looking at the stem of the word and calculating which letter is the most probable of those assigned to the pressed key by using a dictionary.

On pressing a key the order that the associated letters are offered to the user is determined by the likelihood of each associated letter appearing considering the previous letters in the word which are now fixed in position and displayed to the user. If the first character offered is not accepted a further press of the key will offer the second most likely character, this process continues until all possible characters have been offered to the user at which point the process begins again with the most likely character. Since the first statistical database is generated only from the patterns of characters from the start of the string, the statistics will not be affected by patterns of characters that frequently appear in other regions of the string, eg ing which frequently appears at the end, and therefore the accuracy of character selection at the start of the word will increase. The second statistical database only includes words containing more than a defined number of characters e.g. 4 and therefore requires less memory than the corresponding database from EP-A-0842463 which contains words of all lengths.

Since the statistically more likely letters are offered in preference to the less likely letters for a given key, in general, fewer key presses will be required by the user to

type a chosen word. All characters are offered to the
user in turn even if the probability of a particular
character sequence is extremely low. Having selected a
character in a particular position in the string, that
character is then fixed in its position and will not
change regardless of successive characters added to the
string. The algorithm is also adaptable and the
statistical probability of the user using certain words
will be updated taking into account words frequently used
by the user. Words and character patterns will also be
added to the databases in the same way.

The present invention is defined in its various aspects in
the appended claims, to which reference should now be
made.

An embodiment of the invention will now be described in
detail by way of example with reference to the
accompanying drawings in which:

Figure 1 is a typical mobile phone in which several
characters are associated with each key and the selected
characters are displayed on the screen.

Figure 2 is a block diagram showing the process of letter
selection by an embodiment of the invention.

In Figure 1 characters are shown to be associated with the
keys of the phone 10, e.g. The letter J can be presented
on the screen 20 through pressing key 5.

At 210 in Figure 2 a particular key is selected by the
user. The system determines whether the character is to
be positioned within the first 3 characters of the string
at 220. This will be the case for the first three
characters selected. If the letter is positioned within
the first 3 letters of the word the first algorithm is

accessed and the most likely letter associated with the key is presented based on the pattern of letters in the string at 230. At 270 the user determines whether the offered letter is required. If the letter is required the user may proceed and fix the chosen letter to its position within the string at 280. If the presented letter is not required then the user may press the key again and the database will offer the second most likely letter at 260. Once again the user may accept the letter at 270. If the character is still not required then further presses of the key will continue to offer all letters associated with the key.

If the character is not positioned within the first 3 characters of a string at 220 then the second database is accessed at 240. The database searches for the stem of the word in its dictionary. The probability of each letter associated with the key being required is assessed from the word stem and at 250 the most likely character is presented. If the user wishes to accept this character at 290 than he may proceed at 280. If the offered character is not correct then a further press of the same key at 2100 will present the second most likely letter based on statistics. Once again, further presses of the key will bring up successive letters associated with that key until the required letter is presented.

As an example consider a user typing the word HELLO. On an existing mobile phone without predictive text entry, the sequence of key presses is as follows:

4(GHI) 4(GHI) 3(DEF) 3(DEF) 5(JKL) 5(JKL) 5(JKL) 5(JKL)
          H   E   L

5(JKL) 5(JKL) 6(MNO) 6(MNO) 6(MNO)
       L   O

With predictive character selection embodying the present invention it is most likely that the word HELLO can be entered as follows:

4(GHI) 4(GHI) 3(DEF) 5(JKL) 5(JKL) 6(MNO)
        H   E   L   L   O

In this example the number of key presses has been reduced from 13 to 6 and the algorithm can be described as follows:

The user wishes to commence the word with the letter H. On depressing the key associated with the letter H, key 4, the database calculates which of the associated letters is most likely to be required to start a word. The letter G has the highest probability and so is initially offered to the user. Since the letter G is not required, a second depression of the key offers the second highest probability letter, the letter H. Since H is required the user may progress to the next letter. The user now wishes to enter the character E and presses the key 3(DEF). Under the rules of character selection with no predefined sequence two presses of the key would be required to select the letter E. However the database considers the probability of each of the letters associated with the key following the letter H as the second letter in the string. The letter with the highest probability is E and so is offered first to the user. The third letter is obtained in a similar way following the string HE at the start of a word.

On selecting a forth character the dictionary database is accessed. In this case the database looks at the stem of the word, HEL, and calculates the probability of the next letter being a J, K or L based of the number of words in the dictionary that begin, HELJ, HELK and HELL. Since the most probable is HELL, L is offered as the forth letter.

Similarly the database presents the fifth letter with reference to the stem HELL.

In a second example the forth letter of the word BENEFIT may be selected. The character sequence BEN has already been entered. On depressing the key 3(DEF), the word may take one of 3 possibilities; BEBD, BENE or BENF. If the dictionary contains 5 words beginning with BEND, 4 words beginning with BENE and 0 words beginning with BENF, the character D will be offered first, followed by E and finally F.

In a simplified embodiment, only the first algorithm based on statistical probability need be used. However, the database increases significantly in size to accommodate words longer than 4 characters and therefore the second embodiment described above is preferred.

Claims

1.   A method for character selection and display for use
in electronic devices from a reduced keypad in which each
key has a plurality of characters assigned to it
comprising the steps of;
a) detecting a key press
b) selecting one of the characters assigned to the key for
display in dependence on the probability of that character
appearing at that location in a character string;
c) selecting a further character assigned to the key for
display if the first displayed character is not a desired
character for that location;
d) repeating step c) until the desired character is
displayed; and
e) assigning the selected character to that location.

2.   The method of claim 1 wherein if the character is the
first character of the string the probability of the
character being the desired character for that location is
determined from a database containing the number of
recognised strings beginning with that character.

3.   The method of claim 1 or 2 wherein if the character is
positioned within a predefined number of characters from
the start of the string the probability of that character
being the desired character for that location is
determined from a database containing the statistical
probability of that character following the previous
characters in the string.

4.   The method of claims 1, 2 or 3 wherein if the
character is positioned after a predefined number of
characters from the start of the string, the probability
of that character being the desired character for that
location is determined from a dictionary database.

5. The method of claims 2, 3 or 4 wherein said databases are adaptive.

6. The method of claims 1, 2, 3, 4 or 5 wherein, upon pressing a key the order with which the characters associated with the key are displayed to the user is dependent on the probability of the characters being the desired character for that location in the string as determined by said databases.

7. A method for character selection as claimed in claim 1 substantially as herein described, with reference to the accompanying drawings.

8. An apparatus for character selection and display for use in electronic devices from a reduced keypad in which each key has a plurality of characters assigned to it comprising the steps of;
a) detecting a key press
b) selecting one of the characters assigned to the key for display in dependence on the probability of that character appearing at that location in a character string;
c) repeatedly selecting a further character assigned to the key for display if the first displayed character is not a desired character for that location until the desired character is displayed; and
d) assigning the selected character to that location.

9. The apparatus of claim 8 wherein if the character is the first character of the string the probability of the character being the desired character for that location is determined from a database containing the number of recognised strings beginning with that character.

10. The method of claim 8 or 9 wherein if the character is positioned within a predefined number of characters from the start of the string the probability of that

character being the desired character for that location is
determined from a database containing the statistical
probability of that character following the previous
characters in the string.

5   11.   The apparatus of claims 8, 9 or 10 wherein if the
character is positioned after a predefined number of
characters from the start of the string, the probability
of that character being the desired character for that
location is determined from a dictionary database.

10   12.   The apparatus of claims 9, 10 or 11 wherein said
databases are adaptive.

13.   The apparatus of claims 8, 9, 10, 11 or 12 wherein,
upon pressing a key the order with which the characters
associated with the key are displayed to the user is
15   dependent on the probability of the characters being the
desired character for that location in the string as
determined by said databases.

14.   An apparatus for character selection as claimed in
claim 8 substantially as herein described, with reference
20   to the accompanying drawings.

| Application No: | GB 0107931.8 | Examiner: | Melanie Gee |
| Claims searched: | 1-14 | Date of search: | 10 October 2001 |

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

    UK Cl (Ed.S): G4H (HKK, HKV)

    Int Cl (Ed.7): G06F

Other: Online: WPI, EPODOC, PAJ

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
X	EP 0924594 A2 (NOKIA MOBILE PHONES), see whole document.	1, 2, 3, 6, 8, 9, 10, 13
X	US 6005495 A (CONNOLLY et al.), see whole document.	1, 2, 3, 6, 8, 9, 10, 13
X	US 4737980 A (CURTIN et al.), see col. 3 line 38 - col. 4 line 47.	1, 3, 6, 8, 10, 13

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.